Specification Worksheet

version 11/09/2009

REQUEST TYPE (office use only - to be completed by ResDAC)							
CCW:		RDDC:					ResDAC TA:

STUDY/PROJECT INFORMATION

Date:	***June 18, 2010***
Project/Study Name:	**"The Role of Publically Financed Insurance in Massachusetts Health Care Expenditures"**
Linkage to another DUA required:	***Specify DUA number:***
Data Extract Specification:	***2007-2008 MA state residents as identified from 2007-2008 Beneficiary Summary File Records***

REQUESTER INFORMATION

(1) Project Contact (person who will be responsible for operational and acquisition questions about the data)

Name:	**Daniel Gilden**				
Organization:	**JEN Associates, Inc.**				
Address:	**5 Bigelow Street**				
City:	**Cambridge**	**State:**	**MA**	**Zip Code:**	02139
Telephone:	**617-868-5578**			**Fax:**	617-868-7963
Email:	dmg@jen.com				

(2) User (person who signed DUA #16)

Name:	**David Wessman, CIO**				
Organization:	**Massachusetts Division of Health Care Policy and Finance**				
Address:	**2 Boylston Street**				
City:	**Boston**	**State:**	**MA**	**Zip Code:**	02116
Telephone:	**617-988-3272**			**Fax:**	
Email:	david.wessman@state.ma.us				

(3) Custodian (person who signed DUA #17)

Name:	**David Wessman, CIO**				
Organization:	**Massachusetts Division of Health Care Policy and Finance**				
Address:	**2 Boylston Street**				
City:	**Boston**	**State:**	**MA**	**Zip Code:**	02116
Telephone:	**617-988-3272**			**Fax:**	
Email:	david.wessman@state.ma.us				

SHIPPING INFORMATION							
Delivery Service:							
Delivery Service Account Number:							
Special Instructions:							
Ship data to:		**Project Contact**		**User**	x	**Custodian**	

METHOD OF PAYMENT (researcher will be contacted for payment after request is approved)	
Check/Money Order	
Interagency Agreement Number	
Purchase Order (government agencies only)	x

OPERATING SYSTEM (select the operating system that will be used to decrypt and decompress the SDA (*typically delivered on Windows NTFS formatted USB hard drive*). MS Access is not a suitable receiver for data extraction because it does not allow for case sensitivity.	
Windows (Windows 2000 or higher)	x
Unix	
HP-UX 11i or above (PA-RISC only)	
IBM AIX 5.2 or above	
Red Hat Enterprise Linux 3.0 or above (x86 only)	
Solaris 8 or above (SPARC only)	

OUTPUT MEDIA Select ALL acceptable media (BCSSI will determine most appropriate delivery media)	
CD or USB Hard Drive (USB HD may be sent if the data volume is greater than 52 gigs)	x
DVD or USB Hard Drive (USB HD may be sent if the data volume is greater than 250 gigs)	x
USB Hard Drive	x

CMS Disclaimer—User Agreement
Privacy Protected Data—Custom Requests
April 26, 2004

The Center for Medicare & Medicaid Services (CMS) is responsible for administering the Medicare, Medicaid and State Children's Health Insurance Programs. Our agency resources, including staff and computing resources, are primarily dedicated to agency operations. CMS is committed to providing data to other Federal agencies and to the public according to law and as our resources permit. CMS supports these requests with the resources available after agency mission needs have been met.

The increase in CMS mission responsibilities resulting from enactment of the Medicare Drug Improvement and Modernization Act (MMA) has further strained our staffing and computing resources. This disclaimer details the restrictions on CMS services in supporting data requests so that data requestors can plan their projects accordingly. It also specifies the responsibility of the data user and of CMS in regard to the delivery, processing, and understanding of the data files.

Timeframes for data delivery: The HIPAA Privacy Rule, 45 C.F.R. Parts 160 and 164, now require that privacy-protected data requests are approved by the CMS Privacy Board. This board meets monthly to review all requests on their docket. The Board is unable to review every outstanding request within a one month period. Therefore, privacy approval may take up to two months or more. Data processing can take an additional 3-6 months, depending on whether the request is for current or archived files. Therefore, please estimate at least 6-8 months for a request to be processed. CMS will make every effort to process requests in a timely manner, but we cannot guarantee that resources are available to meet any timeframe.

Data accuracy: CMS publishes data that is used by the agency for operational purposes. We use agency standard matching and cross-referencing routines. The requestor accepts the agency data and the agency routines used to produce the data. CMS cannot commit resources to explain or validate its complex matching and cross-referencing programs to requestors.

CMS also publishes the best and most complete documentation available about the file formats and the data. CMS does not insure 100% accuracy of all records and all fields. Some data fields that are not used for agency functions may contain incorrect or incomplete data. Users must familiarize themselves with the detailed data dictionary that is included with every file and published on the internet (http://www.cms.hhs.gov/IdentifiableDataFiles/). A history of each data element, including changes, quality issues, and corrections, is in the data dictionary. Users accept the quality of the data they receive. CMS will not resolve data discrepancies or data questions raised by users. If users would like to report a systemic problem with the data, they may do so. CMS may not have the resources to verify the discrepancy. If the problem is verified, CMS will revise its data documentation.

Data integrity: CMS will ensure that each requestor receives the data requested. Questions about the data must be addressed to CMS within 90 days of receipt. Any alteration of the original data, including conversion to other media or other data formats, is the responsibility of the requestor. Data that has been manipulated or reprocessed by the user is the responsibility of the user. CMS will discuss only the original data delivered to determine that the initial request has been properly processed. CMS has no responsibility for the data after it has been converted, processed or otherwise altered. CMS has no responsibility for assisting users with converting the data to another format.

DATA EXTRACTION DETAIL

[ ] **Finder File provided by Researcher**

[ ] **SSN** - submit 25 byte file with carriage return after each SSN. File format: 1-9 SSN, 14-14 gender, 16-25 DOB mm/dd/yyyy (8 digits with 2 slashes)** See technical publication RDDC-04 on the ResDAC website for details. (http://www.resdac.umn.edu/Tools/TBs/RDDC-04_Finder_File_Encryption_Policy.pdf)

[ ] **HIC** - submit 25 byte file with carriage return after each HIC. File format: 1-12 HIC, 14-14 gender, 16-25 DOB mm/dd/yyyy (8 digits with 2 slashes)** See technical publication RDDC-04 on the ResDAC website for details. (http://www.resdac.umn.edu/Tools/TBs/RDDC-04_Finder_File_Encryption_Policy.pdf)

[ ] **BID (from Acumen) or BENE_ID (from Buccaneer)**

[x] **Other** (may include MSIS_ID, UPIN, Providers, RES_ID/STATE_ID)

*** The information on DOB or gender will only be used if there is a SSN or HIC that is being shared by two persons.*

[ ] **Finder File constructed by BCSSI**

[ ] **BENE_ID Finder File to be constructed using Researcher's provided beneficiary sample criteria** *- data request involves generating a finder file of* ***beneficiaries*** *to be run against the claims and/or enrollment data*

[ ] **Provider Finder File to be constructed using Researcher's provided sample criteria** *- data request involves generating a finder file of* ***providers*** *to be run against the claims data*

[ ] **Data Extract based on standardized percentage selection** *- data request involves extracting claims or enrollment data for a percent sample of beneficiaries*

	Select Files for Extraction: RIF data files will be delivered in a fixed column format with SAS programs (for SAS users) and FTS files (for non-SAS users).			***Indicate percent*** *[assumes enhanced unless stated otherwise]* **5%**	**100%**
CLAIMS/EVENTS					
x	Inpatient	Years	2007-2008		x
x	Outpatient	Years	2007-2008		x
x	SNF	Years	2007-2008		x
x	Hospice	Years	2007-2008		x
x	Home Health	Years	2007-2008		x
x	Carrier	Years	2007-2008		x
x	DMERC	Years	2007-2008		x
x	Part D Event	Years	2007-2008		x
	Drug Characteristics	Years			
x	*Plan Characteristics File*	Years	2007-2008		
MEDPAR					
	All (ss/ls/snf)	Years			
	SS/LS	Years			
	SNF	Years			
ASSESSMENTS*					
	MDS	Years			
	from CMS (based on ___TARGET_DATE or ___SUBMISSION_DATE)				
	from CCW (based on EFFECTIVE_DATE)				
	OASIS (based on ASMT_EFF_DATE)	Years			
	IRF-PAI (based on DSCHRG_DT)	Years			
	Swing Bed (based on EFFECTIVE_DT)	Years			
ENROLLMENT/SUMMARY FILES					
	Denominator File	Years			
	Part D Denominator File	Years			
	Beneficiary Summary File	Years			
x	Beneficiary Summary File with Part D	Years	2007-2008		
	Chronic Condition Summary File	Years			
	Beneficiary Annual Summary File	Years			
	EDB User View	Years	Current		
	(for EDB User View, specify 'most current' or 'all occurrences' on variable selection sheet)				
x	Vital Status File	Years	Current		
x	*Include living beneficiaries*				
x	*Include deceased beneficiaries*				
x	*Include names (requires special permission)*				
x	*Include addresses (requires special permission)*				
MAX DATA					
	All (PS, IP, RX, OT, LT)	Years			
	PS (Personal summary records)	Years			
	IP (Inpatient records)	Years			
	RX(Drug records)	Years			
	OT (Other services records)	Years			
	LT (Long Term Care records)	Years			

**CCW includes only those assessments for which Medicare beneficiaries can be identified.*

SELECT FILES FOR EXTRACTION (cont'd)

MISCELLANEOUS

	MPIER (All)	Years	
	MPIER (Active Only)	Years	
	UPIN Member File	Years	
	Other (Specify)	Years	

CROSSWALKS

	Identifier Crosswalk Buccaneer BENE ID to HIC
x	Identifier Crosswalk Buccaneer BENE ID to SSN
	Identifier Crosswalk Buccaneer BENE ID to MCBS ID
	Identifier Crosswalk Acumen BID to Buccaneer BENE ID
	Identifier Crosswalk Acumen BID to Buccaneer BENE ID to HIC
	Identifier Crosswalk Acumen BID to Buccaneer BENE ID to SSN
	Identifier Crosswalk Acumen BID to SSN to HIC
	MSIS_ID to Buccaneer BENE ID
	MAX BID (Acumen) to Buccaneer BENE ID
	Other (Specify)

CCW CHRONIC CONDITIONS

Note: Finder files containing any personal health or identifying information must be encrypted *prior to* sending to BCSSI.

Select pre-defined chronic condition(s) for cohort and/or control. *(Select all that apply.)*

Cohort				Control			
Include	**Exclude**		**Reference Year(s)***	**Include**	**Exclude**		**Reference Year(s)***
		Stroke/Transient Ischemic Attack				Stroke/Transient Ischemic Attack	
		Rheumatoid Arthritis/ Osteoarthritis				Rheumatoid Arthritis/ Osteoarthritis	
		Prostate Cancer				Prostate Cancer	
		Osteoporosis				Osteoporosis	
		Lung Cancer				Lung Cancer	
		Ischemic Heart Disease				Ischemic Heart Disease	
		Hip/Pelvic Fracture				Hip/Pelvic Fracture	
		Heart Failure				Heart Failure	
		Glaucoma				Glaucoma	
		Female Breast Cancer				Female Breast Cancer	
		Endometrial Cancer				Endometrial Cancer	
		Diabetes				Diabetes	
		Depression				Depression	
		Colorectal Cancer				Colorectal Cancer	
		Chronic Obstructive Pulmonary Disease				Chronic Obstructive Pulmonary Disease	
		Chronic Kidney Disease				Chronic Kidney Disease	
		Cataract				Cataract	
		Atrial Fibrillation				Atrial Fibrillation	
		Alzheimer's Disease and Related Disorders or Senile Dementia				Alzheimer's Disease and Related Disorders or Senile Dementia	
		Alzheimer's Disease				Alzheimer's Disease	
		Acute Myocardial Infarction				Acute Myocardial Infarction	

** Researchers may select reference years for each chronic condition to describe the cohort/control. For example, "Include" with reference years of 1999-current results in a cohort/control of beneficiaries that ever had the chronic condition. While selecting "Exclude" with the same reference years will eliminate any beneficiaries who had the chronic condition from the control/cohort. The researcher may also stipulate individual reference years of interest if desired.*

OPTIONAL: ADDITIONAL CRITERIA

Select additional criteria to subset (restrict) the requested cohort:

Cohort	
Sex: Male _____ Female _____ Age: Age Range _____(yrs) computed as of _____ (date) Date of Death: Describe selection criteria:___	Race: _____White _____Black _____Asian/Pacific Islander _____Hispanic _____North American Native _____Other _____Unknown
Residence[1]: State(s) *(or finder file name)* : ___ *MA* County(ies) *(or finder file name)* : ___ Zip(s) *(or finder file name)* : ___	Criteria to define "Residence" State of residence as of the latest EDB data 2007-2008 State of residence as of the latest claims data _____ Continuous residency in same state for entire time period _____

[1] *Finder files for state codes should be in 2 character FIPS or SSA format. Finder files for county codes should be in 3 character FIPS or SSA format. County finder files must include state codes. Finder files for zip codes should be in 5 or 9 character format.*

Cohort			
Coverage Status	For entire time period (Y or N)	Break in coverage allowed[2] (Y or N)	Comments
Part A _____			
Part B _____			
HMO _____			
State Buy-In _____			
Medicare Status: _____ Aged without ESRD (MSC=10) _____ Aged with ESRD (MSC=11) _____ Disabled without ESRD (MSC=20) _____ Disabled with ESRD (MSC=21) _____ ESRD only (MSC=31) _____ All ESRD			
Comments:			

[2] *Indicates whether the beneficiary must be eligible for coverage for each month of the reference time period or if breaks in the coverage are acceptable*

CMS MAX Data

BENEFICIARY FINDER FILE TO BE CONSTRUCTED USING RESEARCHER SELECTION CRITERIA -
data request involves generating a finder file of beneficiaries to be run against the claims and/or enrollment data

- [ ] Finder to be created only from MAX data.
- [ ] Finder to be created from MAX and Medicare data, combined, depulicated and run against MAX only.
- [ ] Finder to be created from MAX and Medicare data, combined, depulicated and run against MAX and Medicare.

Step 1: Define selection criteria to construct finder file

Eligibility Criteria (applied against Personal Summary (PS) file)
[Note: 'AND' logic is assumed between variables. However, if 'OR' logic is required between variables, specify in the Contact_Request Info spreadsheet.]

A **Recipients** only (1999-2004 PS file field 51; 2005 PS file field 64)
Recipient Indicator equal to any code except "0" [ ]

B Enrollees by **Restricted Benefits Flag** (1999-2004 PS file field 49; 2005 PS file field 55)
Enrollees whose benefits are not restricted for at least one month, code value of 1 or 4 in at least one month [ ]
Enrollees with restricted benefits (specify codes 2, 3, 5, and/or 6) [ ]

C **SMRF Uniform Eligibility Code**-Most Recent (1999-2004 PS file field 21; 2005 PS file field 30)
All codes (will include individuals not enrolled in Medicaid) [ ]
Medicaid enrollees only (All codes ***except "00" and "99")*** [ ]
Eligibility group selection (the following are 1999-2004 values): [ ]
Aged, codes 11,21,31,41,51 [ ]
Disabled, codes:12,22,32,42,52, 3A [ ]
Child, codes:14,16,24,34,44,48,54 [ ]
Adult, codes:15,17,25,35,45,55 [ ]
Specify codes: [ ]

D **Dual Medicaid/Medicare Eligit** (1999-2004 PS file field 7; 2005 PS file field 10)
All Medicaid enrollees (will include all duals, non-duals and unknown dual eligibility status) [ ]
Non-duals (null) [ ]
All Duals (valid HIC) [ ]

Clinical - if more than 10 codes or groups of codes submit on diskette or CD in .csv format
[Note: 'AND' logic is assumed between variables. However, if 'OR' logic is required between variables, specify in the Contact_Request Info spreadsheet.]

E **Type of Service codes NOTE: There is no Type of Service selection for the IP or RX files.**
LT file record selection
All records in the LT [ ]
Nursing Facility records only (LT field 17 equal to 07) [ ] Years to be searched [ ]

OT file record selection
All records in OT file [ ]
All records in OT file ***except*** those for capitation payments. (1999-2004 PS file field 17 = 20, 21, or 22; 2005 PS file field 23 = 20, 21, or 22) [ ] Years to be searched [ ]
Records with only the SMRF TOS code(s) of [ ] Specify codes [ ] Years to be searched [ ]

Step 2: Standard File Selection [Data extract format is text, comma delimited]

[ ] **Output Files with records output based on match to finder file IDs**

Personal Summary	[ ]	Years of interest:	[ ]
Inpatient	[ ]	Years of interest:	[ ]
Other Therapy	[ ]	Years of interest:	[ ]
Long Term Care	[ ]	Years of interest:	[ ]
Prescription Drug	[ ]	Years of interest:	[ ]

[ ] **Output Files with records selected based on match to finder file IDs AND only records with the criteria of interest.**

Personal Summary	[ ]	Years of interest:	[ ]
Inpatient	[ ]	Years of interest:	[ ]
Other Therapy	[ ]	Years of interest:	[ ]
Long Term Care	[ ]	Years of interest:	[ ]
Prescription Drug	[ ]	Years of interest:	[ ]

Prescription Drug Event Data Variable Selection & Justification Table

x' to request	PDE Variable	Reason for Requesting PDE Element (In a few sentences, provide detailed justification for each element.)	Risk of not receiving element (high, medium, low) If risk is high or medium, please provide explanation.
X	Encrypted Part D Event ID	Unique count of this field distinguishes dispense events from adjustments, deletions and credits counts. Pharmacy management measures use annual counts of dispense events as a measure of fills and re-fills. This data provides information on number of contacts with pharmacists and can be used in measures of therapy adherence.	High
X	Encrypted 732 Beneficiary ID	Required for linkage to Part A and B and other payer data for disease state and other utilization. Required for the creation of person histories relating Part D user with Part A and B utilization.	High
X	RX Claim Control Number	Link for multi-record re-submit and adjudication cycles. Helps distinguish real events from administrative records.	High
	Patient Date of Birth (DOB)		
	Patient Gender		
X	RX Service Date	Required for reporting period definitions and for longitudinal analyses of therapy sequence. Key to the creation of person level measures of therapy utilization and adherence.	High
	Paid Date		
X	Service Provider Identifier Qualifier	Necessary for interpretation of Provider Identifier types. May be used to link pharmacies across coding systems. Typology for where Rx dispensing occurs.	Medium
X	Service Provider Identifier	Cluster analyses of patient patterns of care as a function of pharmacy - used to identify like groups of patients. The question of how providers influence drug utilization and adherence can only be addressed by looking at provider-linked patient groups. For example the role of the pharmacy in detecting duplicate therapies and potential drug-drug interactions.	Medium
X	Prescriber Identifier Qualifier	Necessary for interpretation of Provider Identifier types. May be used to link physicians across coding systems. Typology for MDs	

X	Prescriber Identifier	Cluster analyses of patient patterns of care as a function of prescriber – used to identify like groups of patients. The question of how providers influence drug utilization and adherence can only be addressed by looking at prescriber-linked patient groups. For example the prescriber's role in selecting and monitoring recommended therapies can be addressed with this information. Other sources of this information include Part B physician claims.	Medium
X	Prescription/Service Reference Number	The prescription number supports the count of unique prescriptions as opposed to a count of fill-events. This measure helps put in context a therapy measure independent of days-of-supply. This data is a further tool for ensuring that a dispense event can be distinguished from payment system records.	High
X	Product/Service Identifier	Required for the identification of therapy type. The NDC is key for all detailed analyses of utilization, quality of care and payments.	High
X	Encrypted Plan Contract ID	Required for patient cluster analyses based on Plan Code commonality. Helps explain why certain patient variation based on plan affiliation.	High
X	Encrypted Plan Benefit Package ID	Linkage to benefit parameters for comparison to other payer benefit classes	High
X	Compound Code	Required for the identification of therapy type – supplements the NDC information.	High
X	Dispense as Written/Product Selection Code	Used for the analyses of sources of brand/generic distributions. Required for understanding the level of physician and patient resistance to plan prescribing guidelines.	High
X	Quantity Dispensed	Necessary for standardized price analyses – unit cost multiplier. Important for adherence analyses and quality of care measures.	High
X	Days Supply	Necessary for covered days of therapy analyses related to adherence analyses and quality of care measures.	High
X	Fill Number	Used for the identification of new prescriptions. New fill events represent high risk periods for patient adaptation of the therapy. The information can be complemented by person history measures of new therapy.	Medium
	Dispensing Status		
	Drug Coverage Status Code		

X	Adjustment/Deletion Code	Required to understand function of record as a representation of a fill event or an administrative transaction. Key to ensuring clinical events are not confused with administrative records.	High
	Non-Standard Format Code		
	RX Pricing Exception Code		
X	Catastrophic Coverage Code	Necessary for classifying patient history in the context of catastrophic coverage benefit status. A key variable in looking at adequacy state coverage of low-income elderly.	High
X	Gross Drug Cost Below Out-of-Pocket Threshold (GDCB)	Pricing analysis; comparability to other payers, especially Medicaid	High
X	Gross Drug Cost Above Out-of-Pocket Threshold (GDCA)	Pricing analysis; comparability to other payers, especially Medicaid	High
X	Patient Pay Amount	Comparative Costs analyses between payers and financial burden on patients	High
X	Other True Out-of-Pocket (TrOOP) Amount	Comparative Costs analyses between payers and financial burden on patients. A key variable in looking at adequacy state coverage of low-income elderly.	High
X	Low-Income Cost-Sharing Subsidy Amount (LICS)	Important for measuring transitional access issues as patient becomes Medicaid enrolled. A key variable in looking at adequacy state coverage of low-income elderly.	High
X	Patient Liability Reduction due to Other Payer Amount (PLRO)	Important for measuring transitional access issues as patient becomes Medicaid enrolled. A key variable in looking at adequacy state coverage of low-income elderly.	High
X	Covered D Plan Paid Amount (CPP)	Adjustment of patient financial burden analysis. A key variable in looking at adequacy state coverage of low-income elderly.	High
X	Non-covered Plan Paid Amount (NPP)	Comparative Costs analyses between payers	High
X	Gross Drug Cost	Comparative Costs analyses between payers	High
X	Benefit Phase	Confirms analyses of payment distribution as a function of coverage type in a specific period. Important context for understanding potential variations in A and B utilization	High
	Drug Tier		
X	Prior Authorization	Measure of level of physician and patient response to plan prescribing limitations.	High
	Quantity Limits		
X	Step Therapy	Measure of extent of impact of plan cost control measures as accepted by physicians and patients.	High